THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 273, No. 19, Issue of May 8, pp. 11605–11610, 1998

Printed in U.S.A.

Sequences Outside Recognition Sets Are Not Neutral for tRNA Aminoacylation

EVIDENCE FOR NONPERMISSIVE COMBINATIONS OF NUCLEOTIDES IN THE ACCEPTOR STEM OF YEAST $tRNA^{Phe}$*

(Received for publication, October 29, 1997)

Magali Frugier‡, Mark Helm§, Brice Felden¶, Richard Giegé‖, and Catherine Florentz

From Unité Propre de Recherche n° 9002 du CNRS, "Structure des Macromolécules Biologiques et Mécanismes de Reconnaissance", Institut de Biologie Moléculaire et Cellulaire, 15, rue René Descartes, 67084 Strasbourg Cedex, France

Phenylalanine identity of yeast $tRNA^{Phe}$ is governed by five nucleotides including residues A73, G20, and the three anticodon nucleotides (Sampson *et al.*, 1989, *Science* 243, 1363–1366). Analysis of *in vitro* transcripts derived from yeast $tRNA^{Phe}$ and *Escherichia coli* $tRNA^{Ala}$ bearing these recognition elements shows that phenylalanyl-tRNA synthetase is sensitive to additional nucleotides within the acceptor stem. Insertion of G2-C71 has dramatic negative effects in both tRNA frameworks. These effects become compensated by a second-site mutation, the insertion of the wobble G3-U70 pair, which by itself has no effect on phenylalanylation. From a mechanistic point of view, the G2-C71/G3-U70 combination is not a "classical" recognition element since its antideterminant effect is compensated for by a second-site mutation.

This enlarges our understanding of tRNA identity that appears not only to be the outcome of a combination of positive and negative signals forming the so-called recognition/identity set but that is also based on the presence of nonrandom combinations of sequences elsewhere in tRNA. These sequences, we name "permissive elements," are retained by evolution so that they do not hinder aminoacylation. Likely, no nucleotide within a tRNA is of random nature but has been selected so that a tRNA can fulfill all its functions efficiently.

The specificity of transfer RNA aminoacylation is a crucial step in protein synthesis. Investigations during the last years have shown that the aminoacylation identity of a tRNA is linked to the presence of specific sets of signals allowing both discrimination by cognate aminoacyl-tRNA synthetases (aaRSs),[1] the positive elements, and rejection by noncognate synthetases, the negative elements or antideterminants (1–3). The completeness of a set of positive elements has generally been tested by co-transplantation of the corresponding nucleotides into one or several noncognate host tRNAs that acquire the new aminoacylation properties. In several instances, this approach allowed detection of special requirements for the optimal expression of a given aminoacylation identity set within a host tRNA. Thus, minor elements and conformational features were shown to contribute to aminoacylation identities (*e.g.* Refs. 4–8).

Recognition elements required for phenylalanylation of yeast $tRNA^{Phe}$ were defined in the pioneering work of Uhlenbeck and co-workers (9, 10) as a set of five major elements. These elements correspond to G20, G34, A35, A36, and A73, and their competence to confer phenylalanine (Phe) identity was first demonstrated by transplantation into four host tRNAs that all acquired optimal phenylalanylation capacities (10). Nucleotides involved in tertiary interactions were shown not to contribute to identity by a direct effect (11). Alternatively, expression of Phe identity in the yeast $tRNA^{Asp}$ context has revealed that PheRS is sensitive to fine local structural features, such as the D-loop and variable region structures (12). Finally, in neither study based on sequence comparisons of natural or engineered Phe accepting species, nucleotides within the acceptor stem helix were found important for specificity.

In a previous work, we have been able to create a chimeric tRNA, efficiently recognized and aminoacylated at once by three different aminoacyl-tRNA synthetases including yeast PheRS (6), and found that effective phenylalanylation of this tRNA was dependent, among other features, on the sequence of base pair 2–71 within the acceptor stem. Engineering of a tRNA with multiple specificities was based on the synthesis of a chimeric $tRNA^{Asp}$ containing the recognition sets for yeast PheRS (the five residues listed above), for *Escherichia coli* AlaRS (the G3-U70 base pair, *e.g.* Refs. 13 and 14), and for yeast ValRS (A73 and A35; Ref. 15). Notice that the valine identity residues A73 and A35 are common to the Phe recognition set. Simultaneous optimization of alanylation and phenylalanylation efficiencies could be achieved by insertion of specific structural features (the length of the α- and β-domains within the D-loop shaped to 4 and 2 nucleotides and the length of the variable region extended to 5 nucleotides) and mutation of base pair 2–71 in the amino acid acceptor stem from C-G to G-C (6).

Whereas the structural changes introduced were directed by already established yeast PheRS requirements (11, 12), replacement of base pair C2-G71 by G2-C71 in the chimeric $tRNA^{Asp}$ transcript was guided by our present understanding of *E. coli* $tRNA^{Ala}$ identity. Indeed, this base pair is important

* This work was supported in part by funds from the CNRS, Ministère de la Recherche et de l'Enseignement Supérieur (MRES), and Université Louis Pasteur, Strasbourg. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

‡ Supported by a fellowship from Association pour la Recherche contre le Cancer.

§ Supported by a fellowship from the Fonds der Chemischen Industrie and a TMR fellowship from the European Community.

¶ Supported by a fellowship from MRE and present address: Howard Hughes Medical Institute of Human Genetics, Bldg. 553, Salt Lake City, UT 84112.

‖ To whom correspondence should be addressed: IBMC du CNRS, 15, rue René Descartes, 67084 Strasbourg Cedex, France. Tel.: 33 (0)3 88 41 70 59; Fax: 33 (0)3 88 60 22 18; E-mail: giege@ibmc.u-strasbg.fr.

[1] The abbreviations used are: aaRS, aminoacyl-tRNA synthetase; AlaRS, AspRS, PheRS, SerRS, ValRS, alanyl-, aspartyl-, phenylalanyl-, seryl-, and valyl-tRNA synthetases; DTE, diifthioerythritol.

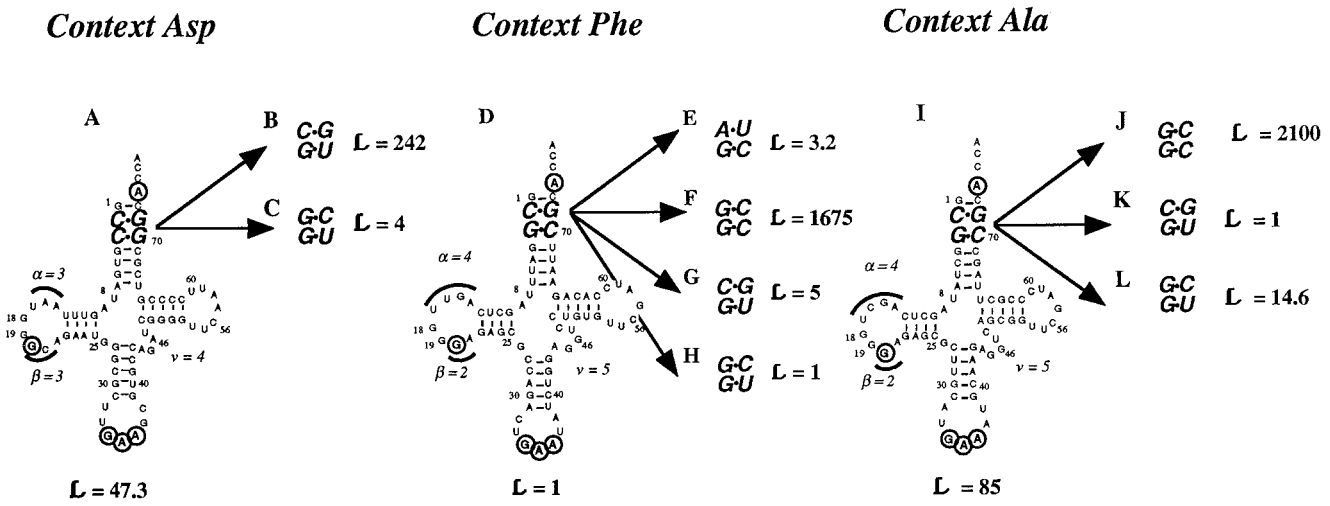

FIG. 1. **Sequence of transcripts derived from yeast tRNA^{Asp} (*A-C*), yeast tRNA^{Phe} (*D-H*), and *E. coli* tRNA^{Ala} (*I-L*).** Sequence data are from Ref. 45 (notice that nucleotide 47 is missing in the variable region of tRNA^{Asp}). The Phe recognition nucleotides are *circled*. Base pairs 2–71 and 3–70, where sequence variations in transcripts occur, are in **bold** characters. To facilitate transcription, all molecules derived from tRNA^{Asp} have G1-C72 instead of U1-A72 in the wild-type tRNA (25). Architectural organization of the D-loop (α- and β-domains that define the location of the conserved G18-G19 dinucleotide in the loop) and length of the variable region are indicated.

for optimal alanine identity expression (5, 16, 17) but was not expected to be of any influence on phenylalanylation (10, 18). Enhancement of phenylalanylation activity by insertion of a G2-C71 base pair was effective in three different structural contexts, all containing the G3-U70 base pair required for efficient alanylation (6).

Here, we basically investigate the role of base pairs 2–71 and 3–70 in Phe identity in yeast tRNA^{Phe} as well as in the frameworks of yeast tRNA^{Asp} and *E. coli* tRNA^{Ala}. The kinetic data collected for various T7 transcripts demonstrate that these base pairs are involved in an optimal expression of the yeast tRNA^{Phe} recognition set. Several possible roles played by these nucleotides are discussed. They do clearly not behave as positive recognition elements as commonly defined in the field. The new outcome of our studies is that the sequence of a tRNA, apart from the recognition elements and the consensus nucleotides involved in the establishment of the three-dimensional structure, is not random for aminoacylation. Some combinations are tolerated, others are not. Thus, specificity is linked to the set of positive and negative recognition elements as well as to an adequate sequence combination within the remaining domains of the tRNA.

EXPERIMENTAL PROCEDURES

Materials—Oligonucleotides were synthesized on an Applied Biosystems 381 DNA synthesizer using the phosphoramidite method and purified by HPLC on a Nucleosyl 125–5-C18 column (Bischoff Chromatography, Zymark-France, Paris). L-[^{3}H] phenylalanine (9.6×10^{11} Bq/mol) was from Amersham France (Les Ulis). Yeast PheRS was a gift of M. Baltzinger (Strasbourg). T7 RNA polymerase was purified according to method as described previously (19). Restriction enzymes (*Bst*N1, *Hin*dIII, and *Bam*HI) and T4 polynucleotide kinase were from New England Biolabs (Beverly, MA). T4 DNA ligase was from Boehringer Mannheim (Meylan, France).

Cloning and in Vitro Transcription—All tRNAs used in this work have been obtained by *in vitro* transcription of synthetic genes. Each of these genes corresponds to the T7 RNA polymerase promoter region directly upstream of the tRNA sequence. The tRNA genes were constructed and cloned into plasmid pUC 119 linearized at *Bam*HI and *Hin*dIII sites according to established methods (20). Tg1 cells were transformed. A *Bst*N1 site coincidental with the 3′-end of the tRNA sequences allows synthesis of tRNAs ending with the expected CCA sequence. Experimental procedures were described previously (20). *In vitro* preparation and purification of transcripts was performed according to established procedures (6). Concentration of stock solutions of transcripts have been determined by absorbency measurements at 260 nm. The primary structure of variants has been checked, in particular for the 5′-end G-rich regions where the sequences were carefully verified by appropriate sequencing methods (21, 22). The global folding of the variant transcripts was shown to be the same as that of wild-type tRNA^{Phe} transcript, as demonstrated by structural mapping of the RNAs by lead according to procedures described previously (23, 24).

Aminoacylation Reactions—Aminoacylation reactions of transcripts derived from tRNA^{Asp} have been performed as described (6) in a medium containing 25 mM Tris-HCl, pH 7.5, 7.5 mM $MgCl_2$, 0.5 mM ATP, 0.1 mg/ml bovine serum albumin, 50 μM [^{3}H]-labeled phenylalanine, and adequate amounts of tRNA transcript and yeast PheRS. Aminoacylation reactions of transcripts derived from tRNA^{Phe} and tRNA^{Ala} have been performed in a medium containing 30 mM HEPES, pH 7.4, 15 mM $MgCl_2$, 12 mM ATP, 30 mM KCl, 4 mM DTE, 50 μM [^{3}H]-labeled phenylalanine, tRNA transcript, and yeast PheRS (10). Before aminoacylation, transcripts were renatured by heating at 65 °C for 90 s and slow cooling to room temperature. Assays were performed in the conventional way (25) with incubation at 30 °C. The kinetic constants were derived from Lineweaver-Burk plots. Since the concentration of amino acids is subsaturating, only apparent kinetic parameters are given. They represent an average of at least two independent experiments. Functional properties of mutants are expressed as catalytic efficiencies of phenylalanylation by k_{cat} over K_m ratios. For easier comparisons, these ratios are also normalized with regard to the wild-type molecule.

RESULTS

Influence of Base Pair G3-U70 on Phenylalanylation within the Yeast tRNA^{Asp} Context—All experiments performed in our previous work (6) leading to the discovery of a positive effect of a G2-C71 base pair on phenylalanylation were done in the tRNA^{Asp} context. These transcripts contained systematically a G3-U70 base pair, necessary for an efficient concomitant alanylation (6). Influence of this last base pair on the chimeric tRNA^{Asp} phenylalanylation activity was tested by comparison with the activity of two types of tRNA^{Asp} transcripts bearing the five primary Phe recognition nucleotides and various combinations at positions 2–71 and 3–70 (Fig. 1). Variant A has the basic sequence of tRNA^{Asp}, the structural framework of tRNA^{Asp}, and the five phenylalanine identity nucleotides (12). Variant B represents the corresponding transcript, containing a G3-U70 base pair. The kinetic parameters of these transcripts (Table I) show that the presence of base pair G3-U70 has only a very limited effect (at most, a 5-fold decrease on aminoacylation efficiency). In this context, when C2-G71 is replaced with G2-C71 (variant C), phenylalanylation efficiency is increased 60-fold. About the same effects were obtained when the structural framework of the host tRNA^{Asp} was engineered to mimic the structural characteristics of tRNA^{Phe} (in

TABLE I

Kinetic parameters for phenylalanylation by yeast PheRS of several transcripts bearing phenylalanine recognition nucleotides

Names of transcripts are according to Fig. 1. Relative k_{cat}/K_m, values are normalized for $(k_{cat}/K_m)_{wt} = 1$ (where wt = wild-type); they vary by 30–50%. **L**-values indicate the loss of catalytic efficiency of a mutant as compared with that of the wild-type tRNA (they correspond to the loss in catalytic efficiency, *i.e.* the inverse of the relative k_{cat}/K_m). Accuracy of K_m and k_{cat} measurements is about 20%.

tRNA transcripts	Base pairs at position 2–71 & 3–70	k_{cat} (s^{-1})	K_m (nM)	k_{cat}/K_m (×1000)	k_{cat}/K_m (relative)	**L** (-fold)
Transcripts derived from yeast tRNA^{Asp}[a]						
D	CG/GC (wt)	2	300	6.66	1	1
A	CG/CG	0.1	717	0.14	0.021	48
B	CG/GU	0.06	2185	0.027	0.004	243
C	GC/GU	2.7	1625	1.66	0.25	4
Transcripts derived from yeast tRNA^{Phe}[b]						
D	CG/GC (wt)	6.3	470	13.4	1	1
E	AU/GC	4.2	1000	4.2	0.31	3
F	GC/GC	0.08	10000	0.008	0.0006	1675
G	CG/GU	8.06	2900	2.8	0.21	5
H	GC/GU	14	1400	10	0.75	1
Transcripts derived from E. coli tRNA^{Ala}[b]						
I	CG/GC (wt)	0.62	4000	0.155	0.01	86
J	GC/GC	0.055	8600	0.0064	0.0005	2096
K	CG/GU	5.75	400	14.4	1.07	1
L	GC/GU	0.8	870	0.92	0.068	15

[a] Phenylalanylation conditions according to Ref. 6.
[b] Phenylalanylation conditions according to Ref. 10.

the D-loop and variable region; Ref. 6). This first set of data led us to suspect the positive effect of the combination of G2-C71/G3-U70 base pairs on phenylalanylation.

Influence of Base Pairs 2–71 and 3–70 on Phenylalanylation within the Yeast tRNA^{Phe} Context—In order to study the role of base pairs 2–71 and 3–70 on phenylalanylation in the cognate natural tRNA framework, a series of mutants of yeast tRNA^{Phe} presenting variations in their sequence have been prepared by *in vitro* transcription (Fig. 1). The wild-type tRNA^{Phe} transcript presents the sequences C2-G71 and G3-C70, and is referred as molecule D in this work (Fig. 1). Four mutants were designed. Variants E and F differ from wild-type tRNA^{Phe} only at the level of the second base pair, where C-G was replaced with A-U (variant E) or G-C (variant F). Variant G differs from wild-type at the third base pair by the presence of a G-U pair. This mutant was constructed to test the role of the alanine identity element G3-U70 alone in phenylalanylation. Finally, the last mutant (variant H) contains two concomitant substitutions in the acceptor stem, namely G2-C71 and G3-U70.

Kinetic parameters characterizing the phenylalanylation capacity of these mutants are summarized in Fig. 1 and Table I. A dramatic effect is observed by inversion of base pair 2–71 from C-G to G-C (variant F), with of more than 1600-fold loss of phenylalanylation efficiency (in comparison to wild-type transcript D). This large decrease is due to a large drop of the rate constant (78-fold) and an increment of K_m by a factor 21 (Table I). Interestingly, the presence of an A-U base pair at this same position (variant E) has no effect on phenylalanylation efficiency.

Conversion of base pair 3–70 from G-C to G-U in the tRNA^{Phe} context (variant G), has only a moderate effect on phenylalanylation (5-fold decrease). The K_m is increased about 6-fold and k_{cat} is not quite affected (Table I). Interestingly, the double mutant (variant H) behaves like the wild-type transcript although it contains a mutation (G2-C71) that affects the aminoacylation in variant F in a strong negative manner. The aminoacylation rate of this mutant is two times better than for the wild-type tRNA, but K_m is increased 3-fold. Thus, as in the tRNA^{Asp} context, the presence of base pairs G2-C71/G3-U70 is positively recognized by PheRS in the tRNA^{Phe} context.

Influence of Base Pairs 2–71 and 3–70 on Phenylalanylation within E. coli tRNA^{Ala} Context—Effect of base pairs 2–71 and 3–70 on phenylalanylation has been further tested within a third tRNA context, namely *E. coli* tRNA^{Ala}. Sequences of the tested transcripts are displayed in Fig. 1. Note that these transcripts share the same fine structural characteristics as tRNA^{Phe} in terms of D-loop and variable region organizations, namely the same length of the α and β regions in the D-loop and the same length of the variable region ($v = 5$). Since constant G18 and G19 residues in the D-loop make long range interactions with the T-loop, in particular the G19-C56 Watson-Crick pair, and that variable region residues are structurally related with D-loop and stem residues (*e.g.* the G15-C48 Levitt pair and the C13-G22-G46 triple), it can be concluded that the core of both tRNA^{Phe} and tRNA^{Ala} transcripts are similar and consequently that the two tRNAs present the same overall three-dimensional structure (1, 6). Variant I, a tRNA^{Ala} with the Phe recognition set and the same sequence as wild-type tRNA^{Phe} at positions 2–71 and 3–70, is 85-fold less well aminoacylated than the reference tRNA^{Phe} transcript. Both k_{cat} and K_m are about 10-fold lower. Interestingly, in variant J, as in tRNA^{Phe}, a G2-C71/G3-C70 combination of nucleotides has a dramatic negative consequence on phenylalanylation with a loss in aminoacylation efficiency of 2100-fold.

Finally, sequence combinations C2-G71/G3-U70 or G2-C71/G3-U70 lead to efficient phenylalanylation of tRNA^{Ala}-derived tRNAs (**L** = 1 and **L** = 14, respectively, for variants K and L), as is the case in the tRNA^{Phe} context (variant H).

RNA Helix Stabilities—Calculation of thermodynamic stabilities of RNA helices was according to Turner *et al.* (26). In the presence of a G3-C70 pair, the free energy of the tRNA^{Phe} accepting helix is between 2 and 3 kcal/mol lower than in the presence of G3-U70. Indeed, in the case of the tRNA^{Phe} framework, the combination G2-C71/G3-C70 confers a free energy of −7.1 kcal/mol, whereas G2-C71/G3-U70 confers only −5.4 kcal/mol. Similarly, in the tRNA^{Ala} framework, a G-C/G-C combination leads to −8.7 kcal/mol, whereas a G-C/G-U combination leads only to −5.7 kcal/mol.

DISCUSSION

General Considerations—First, we recall that the tRNAs investigated here possess the foreseen sequences and fold correctly as verified by sequence analysis and structural mapping with lead (data not shown). Further, these molecules, aside from containing the Phe recognition elements, possess all the same sequences at their amino acid-accepting extremities, namely a G1-C72 base pair extended by a 3′ A73CCA-end. In what follows, and for the sake of simplicity, we assume that the

TABLE II

Ranking of tRNA variants as a function of their efficiency to be phenylalanylated and comparison of the relative contributions of k_{cat} and K_m in charging

Numerical values used for computing normalized kinetic parameters are taken from Table I. According to the definitions given in Table I, $\mathbf{L} = [(K_m)_{mutant}/(k_{cat})_{mutant}]\times[(k_{cat})_{wt}/(K_m)_{wt}]$, which can be rearranged as $\mathbf{L} = [(k_{cat})_{wt}/(k_{cat})_{mutant}] \times [(K_m)_{mutant}/(K_m)_{wt}]$. For a given mutant, k_{cat} and K_m values are normalized for $\mathbf{L} = 1$, so that $\mathbf{L}_N = (k_{cat})_N \times (K_m)_N = 1$. Thus $(k_{cat})_N = [(k_{cat})_{wt}/(k_{cat})_{mutant}]/\mathbf{L}^{1/2}$ and $(K_m)_N = [(K_m)_{mutant}/(K_m)_{wt}]/\mathbf{L}^{1/2}$. For variants with $(k_{cat})_N > 1$, the relative contribution of k_{cat} to **L** is preponderant (these variants have impaired k_{cat} values as compared to wt tRNA); conversely for variants with $(K_m)_N > 1$, the relative contribution of K_m is most important (these variants have a decreased affinity for PheRS as compared to wt tRNA). If for a tRNA variant $(k_{cat})_N \sim (K_m)_N \sim 1$, the relative contribution of both parameters to **L** is the same as for wt tRNA. Normalized k_{cat} and K_m values significantly above 1 are emphasized in bold. For complete sequence of tRNA variants, see Fig. 1. wt, wild-type.

tRNA framework	Name of variants & base pairs 2–71/3–70						**L** (-fold)	Normalized values of k_{cat} and K_m for **L** = 1	
	a C-G G-C	b G-C G-U	c G-C G-C	d C-G G-U	e A-U G-C	f C-G C-G		$(k_{cat})_N$	$(K_m)_N$
Phe		—	—	—	—	—	1	1.00	1.00
Phe	—	H	—	—	—	—	1	0.39	**2.59**
Ala	—	—	—	K	—	—	1	1.14	0.88
Phe	—	—	—	—	E	—	3	0.85	1.19
Asp	—	C	—	—	—	—	4	0.37	**2.70**
Phe	—	—	—	G	—	—	5	0.36	**2.81**
Ala	—	L	—	—	—	—	15	**2.04**	0.49
Asp	—	—	—	—	—	A	48	**2.87**	0.35
Ala	I	—	—	—	—	—	86	1.10	0.92
Asp	—	—	—	B	—	—	243	**2.14**	0.47
Phe	—	—	F	—	—	—	1675	**1.93**	0.51
Ala	—	—	J	—	—	—	2096	**2.50**	0.40

variations in their phenylalanylation properties are essentially due to the nucleotide combinations at positions 2–71 and 3–70, although we are aware that additional effects can occur (see below). To facilitate discussion, data are summarized in Table II. Variants are classified as a function of their decreasing ability to be phenylalanylated (increasing **L**-values) and are divided into six groups (a to f) on the basis of the nature of their 2–71 and 3–70 base pairs, and individual kinetic parameters are normalized so that relative contributions of k_{cat} and K_m to **L** can be easily compared. Several features become immediately apparent. Similar **L**-values can result from different combinations of k_{cat} and K_m. This is for instance, the case when comparing the phenylalanylation capacities of variants H and K. The normalized k_{cat} and K_m show that efficient variants ($\mathbf{L} < 5$) behave phenomenologically either as wild-type tRNA^{Phe} (variants K and E) or are charged in a mechanism where their binding to PheRS is decreased (normalized $K_m > 1$) and k_{cat} improved (variants H, C, and G). Interestingly, in variants with impaired aminoacylation capacity ($\mathbf{L} > 5$), the contribution of k_{cat} is preponderant, except for variant I.

Importance of Base Pair 2–71 and 3–70 for Phenylalanylation in Yeast—Among the 11 tRNA variants possessing the basic Phe recognition set and mutations at base pairs 2–71 and/or 3–70, two mutants with a G2-C71 base pair have drastically decreased aminoacylation efficiencies as compared with wild-type tRNA^{Phe} (variants F and J). The magnitude of the effects, with L-values of 1675 to 2100, is much greater than that observed after mutation of the "classical" Phe recognition elements where losses varied from 10- to 260-fold (10, 18). But, analysis of tRNA variants presenting a G-C/G-U combination at positions 2–71/3–70, shows that the dramatic negative effect brought by G2-C71 in the tRNA^{Phe} and tRNA^{Ala} contexts (variants F and J), is compensated by the G3-U70 base pair (variants H, C, and L). Notably, the G2-C71/G3-U70 combination in the tRNA^{Asp} framework has a positive effect (variant C).

Base Pair C2-G71 Is Not a Positive Recognition Element—According to the current view, tRNA aminoacylations are ensured by a limited number of nucleotides acting as positive signals (1–3). They are presented within optimal structural scaffolds and are generally in direct contact with specific amino acids of synthetases. This definition has two implications. First, mutations at recognition positions should lead to strong losses in aminoacylation efficiency, due to the disappearance of proper hydrogen bonds. Second, transplantation of recognition elements into host tRNA frameworks should be sufficient for acquisition of the corresponding aminoacylation identity by the chimeric tRNAs. The properties of the C2-G71 base pair are not explained by the above scheme. Indeed, our results, combined with those of Uhlenbeck's laboratory, suggest that PheRS makes no distinction between a U-A, A-U, or C-G base pair at position 2–71, and G-C is the only pair to considerably decrease aminoacylation efficiency of the tRNAs chargeable with phenylalanine. Moreover, C2-G71 is not necessary to confer an efficient Phe identity to a host tRNA. Thus, C2-G71 is not a positive identity element.

Alternatively, mutation of a neutral element, by definition an element that does not belong to the recognition set, can also lead to a dramatic loss in aminoacylation efficiency. This is the case if this "neutral" element is replaced by a negative element that can act by introducing a repelling chemical group toward the synthetase or an unfavorable structural context, hindering the correct local or long-range positioning of the substrate. Thus G2-C71 has to be considered as a "negative" element. However, it is not clear if this base pair is in direct contact with the synthetase, although footprinting of natural fully modified tRNA^{Phe} with yeast PheRS showed protection of phosphates 69 and 70 against ethylnitrosourea alkylation (27). The recent crystal structure of the tRNA^{Phe}/PheRS complex from *Thermus thermophilus* (28) indicates that there is no contact between the enzyme and base pairs 2–71 and 3–70 of the tRNA.

Correlated Action of Base Pairs 2–71 and 3–70—Expression of the information carried by base pair G2-C71 is largely dependent on the sequence of base pair 3–70. Its negative effect, observed when position 3–70 is a G-C pair (variants F and J), is balanced by the presence of a G3-U70 pair (variants H and L). Thus, a given base pair affect phenylalanylation differently according to the tRNA framework into which it is embedded. Here again, G3-U70 cannot by itself be considered as a positive recognition signal for yeast PheRS. This conclusion is in line with the great sequence variability of base pair 3–70 among the host tRNAs into which the Phe recognition set has been transplanted (C-G in *S. pombe* tRNA^{Phe}; G-C in *E. coli*, yeast, and wheat germ tRNA^{Phe} and yeast tRNA^{Tyr}; U-A in yeast tRNA^{Met}; and G-U in yeast tRNA^{Arg}) (10).

The positive effect brought by the G2-C71/G3-U70 combination may be related to the structural characteristics of the G-U

pair that decrease the stability of the acceptor stem in comparison with a stem with standard Watson-Crick pairings. We recall that G-U base pairs are important in protein-RNA interactions in general, and in tRNA recognition in particular. For instance, the G3-U70 pair in tRNA^{Ala} is the major alanine identity element for AlaRS and *in E. coli* is involved in a subtle recognition process by the synthetase (4, 29–32). In addition to the presence of the exocyclic amino group in the minor groove of the RNA helix, a G-U pair decreases the helix stability as compared with its G-C equivalent since it contains only two hydrogen bonds (33–35). Unlike a G-C pair, a G-U pair was shown, by NMR, to induce a variation in the local helix geometry (36–38). For example, in an RNA helix mimicking the tRNA^{Ala} acceptor stem, the G-U pair displaces nucleotide C71 and reduces the stacking of the four unpaired nucleotides at the 3′-extremity of the helix (37). The yeast tRNA^{Asp} anticodon helix is another example where a noncanonical G-U pair plays an role in the interaction process with its cognate synthetase by destabilizing an RNA helix. In this case, the G30-U40 pair was shown to be the site of a kink (39) that allows adaptation of the anticodon nucleotides with the AspRS anticodon binding domain (40). Within these lines, we suggest that introduction of a G3-U70 pair into tRNAs inactive for phenylalanylation brings sufficient flexibility into the acceptor stem to allow a fruitful adaptation of the tRNA to PheRS. Similar interpretation has been given to explain the role of G3-U70 in Ala identity *in vivo* (32), although in that case other *in vitro* data argue that direct recognition of this pair by the synthetase is more important than helical distortion (41).

Calculations of the thermodynamic stability (26) of the acceptor helix part estimated for the top four base pairs of the series of variants investigated here argue in this way. The G3-U70 base pair considerably decreases the stability of the acceptor stem, and the flexibility brought by the G2-C71/G3-U70 combination may be sufficient to recover the correct position of the CCA-accepting end. In contrast, when the combination G2-C71/G3-C70 is present, the stability of the acceptor stem is increased, and thus its optimal adaptation on the catalytic domain of the synthetase is prevented. The same conclusion was reached by Saks and Sampson (42), who observed that base pair 3–70 in the acceptor stem of *E. coli* tRNA^{Ser}, which does not contact directly the enzyme, can affect interactions between the neighboring base pairs and SerRS. In this case also, the phenomenon seems to correlate with a decreased flexibility of the acceptor stem, and base pair 3–70 was quoted in this case, a "cryptic" recognition element.

General Conclusions and Perspectives—Involvement of the presently discussed elements at positions 2–70 and 3–71 escaped previous studies although several transplantation experiments were performed. Here, a mutational analysis dictated by serendipitous observations during identity engineering studies (6), has revealed that tRNA nucleotides outside the already reported PheRS recognition set (9, 10) can be of great influence on the expression of Phe identity. This analysis has in particular shown that PheRS accommodates equally well at least three different base pairs at position 2–71. In addition, the present experiments show that PheRS is sensitive to base pair combinations at positions 2–71 and 3–70 in host tRNAs embedding the canonical Phe recognition elements. In particular, the presence of two consecutive G-C pairs at position 2–71 and 3–70 considerably reduces phenylalanylation.

In conclusion, yeast PheRS can accommodate well alternate combinations of base pairs at positions 2–71/3–70 (variants with $L \leq 15$) including a combination where the effect of a G-C base pair can be compensated by a neighboring G-U. This is reminiscent to *in vivo* selection experiments that have shown that a second site mutation can compensate for a defect of an important identity element (43). We propose a fundamentally important structural role of base pairs 2–71 and 3–70 that act as a "permissive box" when G-C/G-U or as a strong "nonpermissive box" when G-C/G-C for PheRS recognition. There is a previous example of a functional box in a tRNA amino acid-accepting domain that contains a G-U wobble pair surrounded by two G-C pairs. In this case, however, the G-U pair is part of an "antideterminant box" accounting for nonrecognition of *E. coli* tRNA^{Sec} by EF-Tu (44).

Considering these facts, it is not excluded that other base pair combinations at positions 2–71 and 3–70 in tRNA^{Phe}, and more generally elsewhere in this structure, present a nonpermissive character or are permissive for PheRS recognition. A systematic search of such combinations will require combinatorial techniques. Altogether, the present findings call for an extended vocabulary for elements involved in tRNA identity. When considering consensus sequences, one has to keep in mind that some nucleotide combinations within "neutral" individual positions might become nonpermissive sequences for some functions of the tRNA. This enlarges tremendously the structural possibility that tRNA-synthetase systems have to ensure their specificity. Moreover, our results suggest that nucleotides within a tRNA have been selected by evolution in such a way that the tRNA can fulfill all its functions in an efficient way. Thus, systematic search of permissive and nonpermissive sequence elements is encouraged, not only in the Phe system, but also in aminoacylation systems of other specificities. Knowledge of such elements would help to understand the relationships between tRNA identities for synthetases and their coupled evolution with that of other identities for other proteins recognizing tRNAs.

Acknowledgments—We thank M. Baltzinger (CNRS/UPR9005, Strasbourg) for the generous gift of yeast PheRS, F. W. Studier (Brookhaven National Laboratory) for T7 RNA polymerase clone, A. Hoeft (IBMC, Strasbourg) for DNA oligonucleotide synthesis, and O. Uhlenbeck (University of Colorado, Boulder) for critical reading of a first version of this manuscript and constructive comments.

REFERENCES

1. Giegé, R., Puglisi, J. D., and Florentz, C. (1993) *Prog. Nucleic Acid Res. Mol. Biol.* **45,** 129–206
2. Saks, M. E., Sampson, J. R., and Abelson, J. N. (1994) *Science* **263,** 191–197
3. McClain, W. H. (1995) in *tRNA: Structure, Biosynthesis, and Function* (Söll, D., and RajBhandary, U. L., eds), pp. 335–347, American Society of Microbiology Press, Washington, D. C.
4. McClain, W., Foss, K., Jenkins, R. A., and Schneider, J. (1991) *Proc. Natl. Acad. Sci. U. S. A.* **88,** 9272–9276
5. Francklyn, C., Musier-Forsyth, K., and Schimmel, P. (1992) *Eur. J. Biochem.* **206,** 315–321
6. Frugier, M., Florentz, C., Schimmel, P., and Giegé, R. (1993) *Biochemistry* **32,** 14053–14061
7. Hou, Y.-M., Westhof, E., and Giegé, R. (1993) *Proc. Natl. Acad. Sci. U. S. A.* **90,** 6776–6780
8. Becker, H. D., Giegé, R., and Kern, D. (1996) *Biochemistry* **35,** 7447–7458
9. Sampson, J. R., and Uhlenbeck, O. C. (1988) *Proc. Natl. Acad. Sci. U. S. A.* **85,** 1033–1037
10. Sampson, J. R., DiRenzo, A. B., Behlen, L. S., and Uhlenbeck, O. C. (1989) *Science* **24,** 1363–1366
11. Sampson, J., DiRenzo, A. B., Behlen, L. S., and Uhlenbeck, O. C. (1990) *Biochemistry* **29,** 2523–2532
12. Perret, V., Florentz, C., Puglisi, J. D., and Giegé, R. (1992) *J. Mol. Biol.* **226,** 323–333
13. Hou, Y.-M., and Schimmel, P. (1988) *Nature* **333,** 140–145
14. McClain, W. H., and Foss, K. (1988) *Science* **240,** 793–796
15. Florentz, C., Dreher, T. W., Rudinger, J., and Giegé, R. (1991) *Eur. J. Biochem.* **195,** 229–234
16. Shi, J. P., Francklyn, C., Hill, K., and Schimmel, P. (1990) *Biochemistry* **29,** 3621–3626
17. McClain, W., Foss, K., Jenkins, R. A., and Schneider, J. (1991) *Proc. Natl. Acad. Sci. U. S. A.* **88,** 6147–6151
18. Sampson, J. R., Behlen, L. S., DiRenzo, A. B., and Uhlenbeck, O. C. (1992) *Biochemistry* **31,** 4164–4167
19. Wyatt, J. R., Chastain, M., and Puglisi, J. D. (1991) *BioTechniques* **11,** 764–769
20. Perret, V., Florentz, C., and Giegé, R. (1990) *FEBS Lett.* **270,** 4–8
21. Brownlee, G. G. (1972) *Determination of Sequences in RNA. Laboratory Techniques in Biochemistry and Molecular Biology* (Work, T. S., and Work, E.,

eds), North Holland Publishing Company, Amsterdam
22. Donnis-Keller, H., Maxam, A. M., and Gilbert, W. (1977) *Nucleic Acids Res.* **4,** 2527–2538
23. Krzyzosiak, W. J., Marciniec, T., Wiewiorowsky, M., Romby, P., Ebel, J.-P., and Giegé, R. (1988) *Biochemistry* **27,** 5771–5777
24. Behlen, L. S., Sampson, J. R., DiRenzo, A. B., and Uhlenbeck, O. C. (1990) *Biochemistry* **29,** 2515–2523
25. Perret, V., Garcia, A., Grosjean, H., Ebel, J.-P., Florentz, C., and Giegé, R. (1990) *Nature* **344,** 787–789
26. Turner, D. H., Sugimoto, N., and Freier, S. M. (1988) *Annu. Rev. Biophys. Biophys. Chem.* **17,** 167–193
27. Romby, P., Moras, D., Bergdoll, M., Dumas, P., Vlassov, V. V., Westhof, E., Ebel, J.- P., and Giegé, R. (1985) *J. Mol. Biol.* **184,** 455–471
28. Goldgur, Y., Mosyak, L., Reshetnikova, L., Ankilova, V., Lavrik, O., Khodyreva, S., and Safro, M. (1997) *Structure* **5,** 59–68
29. Musier-Forsyth, K., Usman, N., Scaringe, S., Doudna, J., Green, R., and Schimmel, P. (1991) *Science* **253,** 784–786
30. Musier-Forsyth, K., and Schimmel, P. (1992) *Nature* **357,** 513–515
31. Musier-Forsyth, K., Shi, J.-P., Henderson, B., Bald, R., Fürste, J. P., Erdmann, V. A., and Schimmel, P. (1995) *J. Am. Chem. Soc.* **117,** 7253–7254
32. Gabriel, K., Schneider, J., and McClain, W. H. (1996) *Science* **271,** 195–197
33. Ladner, J. E., Jack, A., Robertus, J. D., Brown, R. S., Rhodes, D., Clark, B. F. C., and Klug, A. (1975) *Proc. Natl. Acad. Sci. U. S. A.* **72,** 4414–4418
34. Rhodes, D. (1977) *Eur. J. Biochem.* **81,** 91–101
35. Aboul-ela, F., Koh, D., Tinoco, I., Jr., and Martin, F. H. (1985) *Nucleic Acids Res.* **13,** 4811–4824
36. Allain, F. H. T., and Varani, G. (1995) *J. Mol. Biol.* **250,** 333–353
37. Limmer, S., Reif, B., Ott, G., Arnold, L., and Sprinzl, M. (1996) *FEBS Lett.* **385,** 15–20
38. Ramos, A., and Varani, G. (1997) *Nucleic Acids Res.* **25,** 2083–2090
39. Westhof, E., Dumas, P., and Moras, D. (1985) *J. Mol. Biol.* **184,** 119–145
40. Cavarelli, J., Rees, B., Ruff, M., Poterszman, A., Thierry, J.-C., and Moras, D. (1992) in *Structural Tools for the Analysis of Protein-Nucleic Acid Complexes,* pp. 287–298, Birkhäuser Verlag, Basel, Switzerland
41. Beuning, P. J., Yang, F., Schimmel, P., and Musier-Forsyth, K. (1997) *Proc. Natl. Acad. Sci. U. S. A.* **94,** 10150–10154
42. Saks, M. E., and Sampson, J. R. (1996) *EMBO J.* **15,** 2843–2849
43. Hou, Y.-M., and Schimmel, P. (1992) *Biochemistry* **31,** 10310–10314
44. Rudinger, J., Hillenbrandt, R., Sprinzl, M., and Giegé, R. (1996) *EMBO J.* **15,** 650–657
45. Sprinzl, M., Horn, C., Brown, M., Ioudovitch, A., and Steinberg, S. (1998) *Nucleic Acids Res.* **26,** 148–153

Sequences Outside Recognition Sets Are Not Neutral for tRNA Aminoacylation: EVIDENCE FOR NONPERMISSIVE COMBINATIONS OF NUCLEOTIDES IN THE ACCEPTOR STEM OF YEAST tRNAPhe

Magali Frugier, Mark Helm, Brice Felden, Richard Giegé and Catherine Florentz

J. Biol. Chem. 1998, 273:11605-11610.
doi: 10.1074/jbc.273.19.11605